

16th ICCRTS

“Collective C2 in Multinational Civil-Military Operations”

Québec City, Canada, June 21–23, 2011

All-source Information Management and Integration for Improved Collective Intelligence Production

Anne-Claire Boury-Brisset, Anissa Frini, Réjean Lebrun

Defence R&D Canada – Valcartier

2459 Pie-XI North

Quebec, QC, G3J 1X5

Point of contact: Anne-Claire Boury-Brisset

Tel.: (418) 844-4000 #4392

Anne-Claire.Boury-Brisset@drdc-rddc.gc.ca

Paper # 053

Topic #3: Information and Knowledge Exploration

| Report Documentation Page | | Form Approved OMB No. 0704-0188 |
|--|----------------------|---|
| Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. | | |
| 1. REPORT DATE JUN 2011 | 2. REPORT TYPE | 3. DATES COVERED 00-00-2011 to 00-00-2011 |
| 4. TITLE AND SUBTITLE All-source Information Management and Integration for Improved Collective Intelligence Production | | 5a. CONTRACT NUMBER |
| | | 5b. GRANT NUMBER |
| | | 5c. PROGRAM ELEMENT NUMBER |
| 6. AUTHOR(S) | 5d. PROJECT NUMBER | |
| | 5e. TASK NUMBER | |
| | 5f. WORK UNIT NUMBER | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Defence R&D Canada ? Valcartier, 2459 Pie-XI North, Quebec, QC, G3J 1X5, | | 8. PERFORMING ORGANIZATION REPORT NUMBER |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | | 10. SPONSOR/MONITOR'S ACRONYM(S) |
| | | 11. SPONSOR/MONITOR'S REPORT NUMBER(S) |
| 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited | | |
| 13. SUPPLEMENTARY NOTES Presented at the 16th International Command and Control Research and Technology Symposium (ICCRTS 2011), Qu?c City, Qu?c, Canada, June 21-23, 2011. U.S. Government or Federal Rights License. | | |
| 14. ABSTRACT Intelligence analysts face an ever increasing amount of heterogeneous information from various ISR sources, including data from sensors as well as human and open sources provided in disparate multimedia formats and distributed across different systems. Further research is required for enhanced management and integration of all-source intelligence information, and collective production of intelligence products. To this end, a novel approach for global ISR information management and integration is needed. In the paper, we present the vision of a net-centric virtual centralization for the storage and management of all ISR information and intelligence products, making use of standards for data and services. This includes metadata standards for enhanced description of intelligence information for their subsequent retrieval or discovery. Also, semantic technologies and ontologies are envisioned to facilitate ISR integration through mediation of different representation schemes. Initiatives related to geospatial data management are presented. We then derive an architectural framework that incorporates the components and services required for our vision. Such an environment aims at providing users with services to store, access, discover, and integrate intelligence information in support of allsource analysis and sense-making activities. By maximising the use of emerging standards and semantic technologies, the approach should facilitate ISR information integration, collaboration, and interoperability in multinational civil-military operations contexts. | | |
| 15. SUBJECT TERMS | | |

| | | | | | |
|----------------------------------|------------------------------------|-------------------------------------|--|-------------------------------------|------------------------------------|
| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT Same as Report (SAR) | 18. NUMBER OF PAGES 29 | 19a. NAME OF RESPONSIBLE PERSON |
| a. REPORT unclassified | b. ABSTRACT unclassified | c. THIS PAGE unclassified | | | |

All-source Information Management and Integration for Improved Collective Intelligence Production

Anne-Claire Boury-Brisset, Anissa Frini, Réjean Lebrun

Defence R&D Canada – Valcartier

2459 Pie-XI North

Quebec, QC, G3J 1X5

Tel.: (418) 844-4000 #4392

Anne-Claire.Boury-Brisset@drdc-rddc.gc.ca

Abstract

Intelligence analysts face an ever increasing amount of heterogeneous information from various ISR sources, including data from sensors as well as human and open sources, provided in disparate multimedia formats and distributed across different systems. Further research is required for enhanced management and integration of all-source intelligence information, and collective production of intelligence products. To this end, a novel approach for global ISR information management and integration is needed. In the paper, we present the vision of a net-centric virtual centralization for the storage and management of all ISR information and intelligence products, making use of standards for data and services. This includes metadata standards for enhanced description of intelligence information for their subsequent retrieval or discovery. Also, semantic technologies and ontologies are envisioned to facilitate ISR integration through mediation of different representation schemes. Initiatives related to geospatial data management are presented. We then derive an architectural framework that incorporates the components and services required for our vision. Such an environment aims at providing users with services to store, access, discover, and integrate intelligence information in support of all-source analysis and sense-making activities. By maximising the use of emerging standards and semantic technologies, the approach should facilitate ISR information integration, collaboration, and interoperability in multinational civil-military operations contexts.

Keywords: information management, standards, metadata, intelligence.

1. Introduction

With a variety of Intelligence, Surveillance, Reconnaissance (ISR) collection capabilities made available, military operators and intelligence analysts have to deal with ever-increasing amounts of ISR data and information from various sources (SIGINT, IMINT, GEOINT, HUMINT, OSINT, etc.), produced in disparate multiple media formats (raw data sets from sensors, e.g., video, images, sound files, as well as human reports and open source text) and distributed across different systems. These sources are distributed, and data/information is often stored in stove-piped systems. From the user perspective, this

situation results in poor data visibility, and inadequate capability to effectively access, discover and share information sources or intelligence products regardless of their location, to name a few.

Moreover, intelligence analysts should spend more time developing knowledge and intelligence rather than searching or gathering data and information. Still, further research is required to improve the management and integration of intelligence information from multiple sources, for the collective production of intelligence products. This is the focus of an applied research project conducted within the Intelligence and Information Section at Defence R&D Canada (DRDC) – Valcartier. This project aims at investigating advanced concepts, techniques and technologies in order to provide enhanced capabilities for the management and integration of multiple information sources and intelligence products made available to intelligence operators and officers in support of the production of intelligence and sense-making activities. In this paper, we present the vision of a net-centric virtual centralization for the storage and management of multiple ISR information and intelligence products, making use of standards for data and services. Concepts and technologies including automatic metadata tagging, semantic knowledge extraction, multimedia indexing, seamless access, knowledge discovery, smart filtering are being investigated in order to provide the Army and Intelligence Communities with novel capabilities in terms of all source information management and integration.

The paper is structured as follows. In the next section, we describe the intelligence cycle, its intelligence sources and products. Then we introduce the role of information management and the benefits of using standards in support of information management and integration. We next present ongoing developments for geospatial data management. Following is a discussion about the challenges for the integration of intelligence information, and the vision of a service-based information management and integration framework using open and flexible SOA architecture. Aspects related to interoperability and collaboration in this context are finally introduced, and some conclusions and future work described.

2. The intelligence cycle, its sources and products

According to the Canadian Joint Intelligence Doctrine [9] and the Land Force Intelligence Field Manual [8], the intelligence cycle is composed of four steps: direction, collection, processing, and dissemination. More precisely:

- **Direction** consists of determining the intelligence requirements, planning the collection effort, issuing orders and requests to collection agencies and maintaining a continuous check on the productivity of such agencies.
- **Collection** is the process during which information and intelligence are collected from sources and agencies in order to meet the intelligence requirements.
- **Processing** regroups a series of actions which consists of collation; evaluation; analysis and integration; and interpretation of information and/or other intelligence.

- **Dissemination** is the delivery of intelligence and is defined as “The timely conveyance of intelligence, in an appropriate form and by any suitable means, to those who need it”.

Intelligence products are derived from the collection, analysis and integration, and interpretation of tactical data and information from various sources, considered individually as single sources in a first stage, then combined from an all-source perspective. The main intelligence sources, also called collection disciplines include:

- Human Intelligence (HUMINT)
- Imagery Intelligence (IMINT)
- Geospatial Intelligence (GEOINT)
- Measurement and Signature Intelligence (MASINT)
- Signal Intelligence (SIGINT)
 - Communications Intelligence (COMINT)
 - Electronic Intelligence (ELINT)
- Open Source Intelligence (OSINT)
- Technical Intelligence (TECHINT)

These intelligence disciplines produce intelligence products by collecting, processing and analyzing data/information from multiple sources available in different formats and data stores, often stove-piped.

According to the US Army doctrine for intelligence FM-2.0 [6], All-source intelligence is defined as *the products, organizations, and activities that **incorporate all sources of information and intelligence**, most frequently including human resources intelligence, imagery intelligence, measurement and signature intelligence, signals intelligence, and open-source data, in the production of intelligence. All-source intelligence is both a separate intelligence discipline and the name of the process used to produce intelligence from multiple intelligence or information sources.*

In this paper, we use the terms All-source or multi-INT with the same meaning, as we look at the problem of managing and integrating information from multiple intelligence disciplines represented using different schemes, be they information sources or intelligence products.

Simply speaking, intelligence information used in these disciplines falls into two broad categories: *structured information* and *unstructured information*. Structured information conforms to structures defined for a domain, while unstructured information, e.g. collected from open source or contained in human reports, has no underlying model that exposes the semantics of the data. For the latter, some operations are required to make the information automatically exploitable, e.g. by metadata tagging, to add explicit semantics to the data content, or by extracting relevant content from the text (entities, location, etc.).

Related research has been undertaken recently for the integration or fusion of hard/soft information within the information fusion community from different perspectives [7,14].

Hard information refers to data acquired from physical sensors that are structured and can be characterized mathematically, while soft data refers to data based on human observations reported in unstructured/semi-structured format. Enabling techniques for the integration/fusion of hard/soft information proposed by this community are relevant for our work and some are mentioned below.

The domain of data integration from heterogeneous sources has been extensively studied in the literature for years, either by using schema mapping mechanisms, or more recently by exploiting the data underlying semantics through ontological alignment. There are also major initiatives ongoing in the military ISR community to advance the sharing and exploitation of ISR data, for example the development of the Distributed Common Ground/Surface System (DCGS). Moreover, software tools and research prototypes are being developed in support of information management (IM) or information exploitation (IX) for the production of intelligence from multiple sources, e.g. [4, 15].

The contribution of the paper aims at highlighting some advances and challenges for all-source or multi-INT information management and integration, in particular:

- The role of information management in support of the Intelligence cycle,
- The exploitation of standards for IM/IX,
- Ongoing developments for geospatial data management,
- The envisioned challenges for intelligence information integration,
- The vision of a service-based information management and integration framework, and the use of open and flexible SOA architectures.

These are described below.

3. Information management and Standards

Information management

In a general sense, information management refers to the people, processes and technology dedicated to the gathering, organizing, managing, disseminating, leveraging and disposing of all information assets, from raw data to derived products, for use by a community. In the context of the intelligence process described above, effective management of information and intelligence throughout their life cycle is required in order to enable the discovery, access, integration, dissemination, exploitation of actionable intelligence from multiple federated sources by various information consumers and producers, and to facilitate interoperability. Effective information management is essential due to the vast amount of information and the different types of content manipulated by information producers and consumers, and the related challenges of discovering and integrating information/intelligence in this context.

First, as part of an enterprise architecture vision, an information architecture is required to describe the information assets to be managed and how they relate to processes and needs. Second, there is a need for provision of a collection of data/information services to make the various resources accessible, discoverable, and sharable within a virtual information space. The latter must accommodate multiple data formats (structured,

unstructured, streaming, etc.) and various service interaction protocols. As each intelligence information source has its own characteristics in terms of format and structure, it must be described to expose its structure and content and make it discoverable and sharable in a standardized way.

Next we describe the use of metadata standards as a key component to manage heterogeneous information sources and products. We then describe the challenges for the integration of all-source intelligence.

Exploitation of standards

The use of standards is critical to describe information assets in a uniform manner, in order to make information discoverable, to enable enhanced access and exploitation of intelligence information as well as improved interoperability. For that purpose, several organizations such as the International Standards Organization (ISO), the North Atlantic Treaty Organization (NATO), and the Open Geospatial Consortium (OGC), contribute to the advance and development of open standards for enhanced interoperability and management of military information resources, in particular through the development of various metadata standards. Such standardized metadata are being defined by various communities of interest in order to prescribe the characterization of specific properties of their corresponding data resources. For example, there is an increasing interest in full motion video (FMV) in the military and intelligence communities for their analysis and exploitation, as FMV represents observation over time and space that can provide insight about a situation. However, search in large volumes of motion imagery and video based on specific context must be facilitated using appropriate metadata schemes. To this end, metadata standards have to be carefully represented to meet the requirements of intelligence analysts.

Different metadata types can be distinguished to characterize information objects, e.g. descriptive information about the information content (the subject, what it is about, who, where, i.e. semantic elements), context information about the object creation (title, creator, date, location, etc.), or administrative information related to access control, versioning, or traceability.

The use of such standards within their respective information community, and eventually their specialization using profiles to meet specific requirements, make interoperability possible. XML and RDF (Resource Description Framework) are commonly used to represent and store metadata for subsequent search and analysis. Metadata registries associated with data/information resources constitute a foundational component of an effective information management framework.

Below are a set of standards for describing geography and imagery information:

- ISO 19115: geographic information metadata and extensions to imagery metadata and gridded data (ISO 19115-2)
- NATO STANAG 4545 – Secondary Imagery Format (NSIF)

- NATO STANAG 4607 – Ground Moving Target Indicator Format (GMTIF)
- NATO STANAG 4609 – Digital Motion Imagery
- NGA Motion Imagery Standard Profile (MISP) specifications: describe data format and metadata for motion video data
- ISO 19130 – SensorML : description of sensor parameters
- OGC Geography Markup Language (GML) – ISO 19136
- OGC Keyhole Markup Language (KML) – geographic visualization.

Unstructured information management

As mentioned in section 2, unstructured information is common in the intelligence domain (e.g. human reports, text transcripts, web sites). As unstructured information has no underlying structure to facilitate automated interpretation by machines, natural language techniques and tools are necessary to make sense of the semantics within the text. Commercial and open source tools have been developed to support text processing (e.g. information extraction, annotation, categorization) based on statistics or text analysis techniques. In particular, the use of semantic technologies such as domain ontologies are of interest in this context to provide a standardized way to tag or annotate unstructured information. Providing metadata tags that conform to a domain ontology is critical from the perspective of multi-source or multi-INT information integration.

4. Geospatial data management

Geospatial information is a key component for multi-int production, as soldiers, commanders and analysts rely on geospatial information to assess situations and make decisions. For effective exploitation of this type of information, the use of metadata standards is critical to making geospatial information discoverable and retrievable.

DRDC has been highly involved over the last years in improving the management and dissemination functions of geospatial products within the Canadian Forces as part of research activities. The Canadian Mapping and Charting Establishment (MCE) has the responsibility to provide timely and accurate geospatial information and geomatics support to the Canadian Forces, the Department of National Defence and Other Government Departments in accordance with Canada's national objectives. MCE has got several data centres, each having the responsibility to deliver either nautical charts to the Canadian naval fleet, flight maps to the aeronautical fleet or topographic maps and imagery to soldiers on the ground. Together, these maps represent more than a million documents which previously existed in silos and with heterogeneous description formats, thus complicating the task to find and assemble data for a given mission in a timely fashion. To overcome this problem, MCE has been putting together the Gateway Services (GWS), aimed at modernizing and centralizing the geospatial data management and dissemination functions.

As part of research and development projects, DRDC has designed and developed two key components of GWS, namely GeOLAP (Geospatial Online Analytical Processing)

and GeoMIS (Geospatial Metadata and Inventory System). GeOLAP provides a web-based infrastructure to facilitate the online discovery, the ordering and exploitation of geospatial data. GeoMIS is the back-bone application which aims at managing centrally, through the use of metadata standards (ISO 19115, ISO 19139), the distributed inventories of geospatial products within MCE while providing advanced ordering and client management functionalities [10]. To facilitate metadata management, automated metadata extractors have been developed for numerous geospatial data products formats. Search over the collection of geospatial products is offered to the user based on various spatial criteria, and various filters exploiting the metadata set.

Moreover, the GeoMIS application has the ability to export its metadata content in an ISO 19139 XML implementation, as well as to export map footprints to various formats including KML, thus making possible further exploitation and visualization in Google Earth.

These applications have greatly simplified the management of such large geospatial inventories and have reduced from days to hours and minutes the time required to assemble and to prepare geospatial data in support to the Canadian Forces missions. These two components of the GWS are now fully deployed on the intranet network of the Department of National Defence.

For the data held in GeoMIS to be exposed in a service-oriented architecture, a service-based access to the application is provided by using the Open Geospatial Consortium (OGC) Catalog Service for the Web (CSW) specification. As GeoMIS actually complies with the ISO 19115 metadata specification and has the ability to export its metadata content in an ISO 19139 XML implementation, it can expose its metadata within a CSW implementation which can consume ISO 19139 XML documents. The CSW client provides advanced search and filtering capabilities against the server. The complete geospatial inventory can now be exposed to any allied countries through standardized OGC-compliant catalog.

Beyond the geospatial products being managed for MCE, research investigations are conducted to extend the application to the management of additional imagery data from multiple sensors (hyperspectral, multispectral, panchromatic, LiDAR, RADAR). In order to exploit complex forms of geospatial data such as hyperspectral imagery, there is a need to describe the products data in greater details. In this context, we are exploring the opportunity of using the ISO-19115-2 metadata standard extension for gridded data, to assess its suitability to describe and subsequently discover this kind of complex imagery data.

5. All-Source information integration

In the intelligence context, integration of intelligence information from multiple sources for the production of intelligence involves the reconciliation of unstructured and structured worlds, e.g. information from sensors, with intelligence reports from human observers, so that intelligence analysts, supported by tools, can access data that are relevant to support sense-making activities, and can easily establish links among these data. While structured data conform to a representation structure that facilitates their data

access, unstructured data must be enriched with semantic metadata elements, in order to facilitate their discovery and enable subsequent link analysis.

While several metadata standards being developed are of interest for the intelligence community, some challenges remain for the integration of these metadata standards within a federated environment. The OGC has just started a “Fusion Standards Study” [13] with the objective to examine the relevant standards and challenges for the fusion of multi-INT sources. In particular, we can anticipate that metadata mapping across domains, i.e. communities of interest (COIs) will necessitate the use of translators/converters to map the metadata elements between domains, supported by some domain taxonomy or ontology.

In the literature, Pravia and colleagues [14] proposed mechanisms to correlate hard and soft data, in particular through the annotation of soft data to enable association with the hard information elements. This is illustrated with the use of a real data set. Also, Solano and Tanik [16] proposed a framework for combining multi-source metadata with predictive and analytic algorithms to synthesize actionable knowledge (integration of the JDL fusion levels 2 to 4). For that purpose, they provide an extension to the ISO 19115 model in order to capture the necessary metadata for their fusion framework.

In a recent study for information management investigating the use of metadata in a multi-source environment, including multiple communities with their own terminologies, Veres and Ng [17] propose a solution making use of RDF for representing metadata, and exploiting the SUMO upper-ontology to map the various metadata schemes and their respective underlying domain ontologies. As a starting point, they have initiated the mapping of the US DDMS (Defence Discovery Metadata Specification) to SUMO.

While our project is at a preliminary stage from a multi-int perspective, we think that metadata, enriched with explicit semantics described in ontologies (e.g. semantic metadata elements described in OWL) would provide the required semantic foundation to enhance data access and discovery services in a multi-int environment, as well as the building of more advanced reasoning services. To this end, a reference ontology has to be selected. The JC3IEDM model or the Universal Core (UCore) would be worth examining in this context.

6. Information management architectural framework

The main objective for our research is to provide intelligence analysts with an information-centric environment facilitating the management and integration of a large amount of distributed heterogeneous multimedia data/information sources in support of the production of intelligence and sense-making activities. The derived architectural framework is a service-based information management environment incorporating services that facilitate access to data and discovery, integration of data from diverse sources, handling of large volume of data, and providing analytical services (e.g. data mining). Figure 1 presents the high-level architectural framework for All-source

information integration and management (ASIM). This framework aims at providing the necessary information integration and management services that will be utilized in support of various information exploitation services, such as analytical services (represented at the top on the figure).

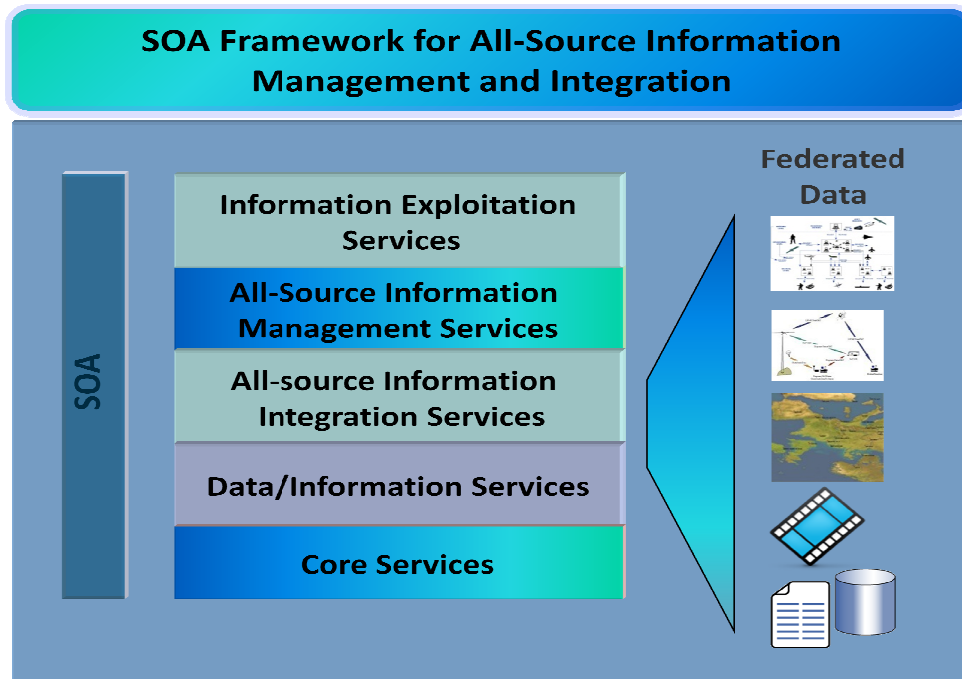


Figure 1: ASIM Framework

We next present the benefits of a service-oriented architecture, and then describe the different categories of services that are envisioned.

Service oriented architectures

Service Oriented Architecture (SOA) has emerged as the predominant paradigm for the building of flexible and scalable architectures in net-centric environments. SOA is an architectural discipline that relies upon the exposure of a collection of loosely-coupled, distributed services which communicate and interoperate via agreed standards across the network. Some benefits are directly based on the principles of service orientation, mainly: services are loosely coupled, autonomous, discoverable, composable and reusable.

The services can be composed into higher-level applications to support agile business processes. Consequently, it is desired to provision for information services that can be deployed within a SOA environment to support intelligence functional processes needed by users.

Categories of services

Based on the information management approach and services envisioned in our environment, we can derive a framework organized around different categories of services in order to make the various distributed datasets available to the federation of information management (IM) and information exploitation (IX) services to intelligence producers and consumers: these include core services, data/information services, information integration/management services, and reasoning services in support of intelligence processes.

Relevant work in this domain provides valuable insight to help carry out the services decomposition. In particular, as part of the NATO Network Enabled Capability (NNEC) initiative, a set of Core Enterprise Services has been consolidated [11]. From another perspective, S. Bray [3] proposed a framework for Warfighter Information Services (WISF), based on the vision of distributed information services and distributed knowledge stores. The framework details and categorizes a rich set of services ranging from information extraction, information collation, to decision support and collaboration, as well as supporting services necessary in a service-oriented architecture. This set of information services is further decomposed and can also be considered as a starting point to articulate our framework.

Compared to Bray's WISF vision and decomposition in categories of information services, the envisioned set of services in our framework is centered around intelligence processes support. Consequently, this set of services (cf. Figure 2), some of which are already being developed and integrated within DRDC's Intelligence Science & Technology Integration Platform (ISTIP), will include the following categories:

- Core SOA services, i.e. enabling SOA (e.g. messaging, discovery, authentication, orchestration, choreography), e.g. based on NATO core enterprise services framework;
- Data/information services: ETL (extract/transform/load), language translation, speech-to-text, unstructured data services (e.g. semantic tagging, annotation, named-entity extraction); data/information access, etc.
- All-Source information/intelligence integration and management services: storage/data insertion, ontology/metadata management, data/information mapping, search and retrieval (semantic, faceted, federated, geospatial), notification/alert, publish, text summarizing/data synthesis, geo-referencing, etc.
- Information/intelligence exploitation services: reasoning/inferencing service, e.g. anomaly detection, geospatial analysis, link analysis, pattern matching, etc.



Figure 2: Categories of services

7. Information sharing and collaboration

Facilitating ISR information sharing and collaboration presents several challenges including governance, as described in [1]. In the following, it is being addressed from different perspectives 1) within a multi-INT environment in order to promote sharing of intelligence information and collaboration among intelligence producers and consumers, and 2) in the context of multinational civil-military coalition operations where standards and policies are critical enablers.

As mentioned above, intelligence information can be shared, integrated and exploited by information consumers if it is made visible and discoverable. Thus, the collective production of intelligence from multiple intelligence sources within a distributed environment will be made possible through the use of multiple metadata standards. These should be enriched with taxonomies/ontologies capturing their underlying semantics to facilitate cross-domain information sharing among different intelligence disciplines. While different approaches can be adopted for the collaborative production of descriptive semantic metadata by intelligence analysts, 1) user-defined tags (folksonomy), or 2) tags based on a controlled vocabulary or domain taxonomies, the latter approach should facilitate interoperability and collaboration based on a common semantics, although it may necessitate the use of translators across domains.

In today's multinational military-civil operations, information sharing must be enabled through the use of standards and well-defined protocols. With the objective to enforce the interoperability of ISR systems, the NATO Multi-Sensor Aerospace-ground Joint ISR Interoperability Coalition (MAJIIC) project developed standards, technologies and an architecture. The key component of the architecture is the Coalition Shared Data (CSD) server based on NATO standardization agreements such as STANAG 4559 for imagery, 4609 (motion imagery), 4607 (GMTI) and 5516 (tracks), that makes use of a synchronization mechanism of the metadata between the different servers [5]. The use of these standardized metadata facilitates the storage and retrieval of ISR data by users interacting with the system. MAJIIC is now entering a new phase (MAJIIC 2) where additional research areas will be investigated, including cross COI interoperability, multi-INT and multi-source analysis, which are of great interest in our context. Consequently, we will align our efforts with this international initiative to contribute and benefit from the advances in this area.

8. Conclusions and perspectives

In this paper, we presented a research initiative undertaken towards the development of a net-centric information environment for the storage, management and integration of intelligence information and products, making use of standards for data and services. This includes metadata standards for enhanced description of intelligence information for their subsequent retrieval or discovery. Also, semantic technologies and ontologies are considered to facilitate information integration through the mediation of different representation schemes. The envisioned multi-INT environment should benefit from lessons learned from the geospatial data management initiative, while leveraging from efforts underway for metadata harmonization, and being aligned with MAJIIC phase-2 research efforts.

By maximising the use of emerging standards and semantic technologies, the approach should facilitate ISR information integration, collaboration, and interoperability in multinational civil-military operations contexts.

From a technological perspective, we should conduct further research in line with NGA research directions [12], in particular research on database technology and spatial data infrastructures to handle multi-scale, multi-dimensional, multi-temporal data, as the database requirements for extremely high spectral and spatial resolution, multimedia imagery and free form text will continue to challenge most existing data schemas and models. The use of data clouds will also be investigated as a potential starting point for the connection of data silos, given the required scalability and the number of data schemas and standards that the architecture take into account.

Other topics to be investigated further include security policies (access control), and user-centricity as we aim at providing users with information and services that are relevant in support of their intelligence processes according to their roles and privileges [2].

References

- [1] D. M. D'Agostino, Intelligence, Surveillance, Reconnaissance, Overarching Guidance is Needed to Advance Information Sharing, US Government Accountability Office, GAO-10-500T, March 2010.
- [2] A. Badia, Personal information management for intelligence tasks, Intelligence and Security Informatics, 2008.
- [3] S. Bray, A Framework for Warfighter Information Services – using the Concept of a Virtual Knowledge Base, Proceedings of the 15th International Command and Control Research and Technology Symposium (ICCRTS), 2010.
- [4] Brynielsson J., Horndahl A., Kaati L., Martenson C., Svenson P., Development of Computerized Support Tools for Intelligence Work, in Proceedings of the 14th International Command and Control Research and Technology Symposium (ICCRTS), 2009.
- [5] B. Essendorfer, W. Mueller, Interoperable sharing of Data with the Coalition Shared Data Server, Proceedings of NATO RTO IST-086 Symposium, Crisis Management, Bucharest, Roumania, May 2009.
- [6] FM 2-0, “Intelligence”, Headquarters Department of the Army, Washington, DC, 23 March 2010.
- [7] D.L. Hall, J. Llinas, M. McNeese, T. Mullen, A Framework for Dynamic Hard/Soft Fusion, Proceedings of the 11th conference on Information Fusion, Cologne, Germany, 2008.
- [8] Intelligence field manual, Department of National Defence, land forces information operations, B-GL-357-001/FP-001, 30 Jan 2001.
- [9] Joint Intelligence Doctrine, B-GJ-005-200/FP-000, 2003-05-21. Canada, National Defence, mai 2003.
- [10] Lebrun, R., Rowlands, D., Gateway Services: Automating Metadata Capture and Dissemination of Geospatial Products Using ISO 19115, GEOINT 2010 Symposium, New Orleans, November 2010.
- [11] NATO NNEC Core Enterprise Services, NATO NC3I, December 2009.
- [12] NGA, New Research Directions for the National Geospatial-Intelligence Agency, Steering Committee on New Research Directions for the National Geospatial-Intelligence Agency; Mapping Science Committee; National Research Council, National Academy Press, 70 pages, 2010, <http://www.nap.edu/catalog/12964.html>, (last accessed January 2011).
- [13] OGC, Open Geospatial Consortium, <http://www.opengeospatial.org/projects/initiatives/fusion2>, (last accessed January 2011).
- [14] M. A. Pravia, O. Babko-Malaya, M. Schneider, J. V. White, C-Y Chong, A.S. Willsky, Lessons Learned in the creation of a dataset for Hard/Soft Information Fusion, Proceedings of the 12th International conference on Information Fusion, Seattle, USA, 2009.

- [15] Roy, J. and Auger, A., The Multi-Intelligence Tools Suite – Supporting Research and Development in Information and Knowledge Exploitation, 16th International Command and Control Research and Technology Symposium (ICCRTS), “Collective C2 in Multinational Civil-Military Operations”, Québec City, Canada, June 21–23, 2011.
- [16] M. Solano, M. Tanik, Recombinant Cognition Synthesis – An introduction, Proceedings of the 11th International Conference on Information Fusion, Cologne, Germany, 2008.
- [17] C. Veres, S. Ng, An Approach to Information Management for AIR7000 with Metadata and Ontologies, DSTO-TR-2289, DSTO, Oct. 2009.

Defence Research and Development Canada

All-source Information Management and Integration for Improved Collective Intelligence Production

Anne-Claire Boury-Brisset
Anissa Frini
Réjean Lebrun

Intelligence and Information Section

16th ICCRTS
June 2011



Outline

- Context
- All-Source Intelligence
- All-Source Information Management challenges
- Geospatial Data Management
- Service-based Information Management and Integration Framework
- Conclusions and Perspectives

Context and objectives

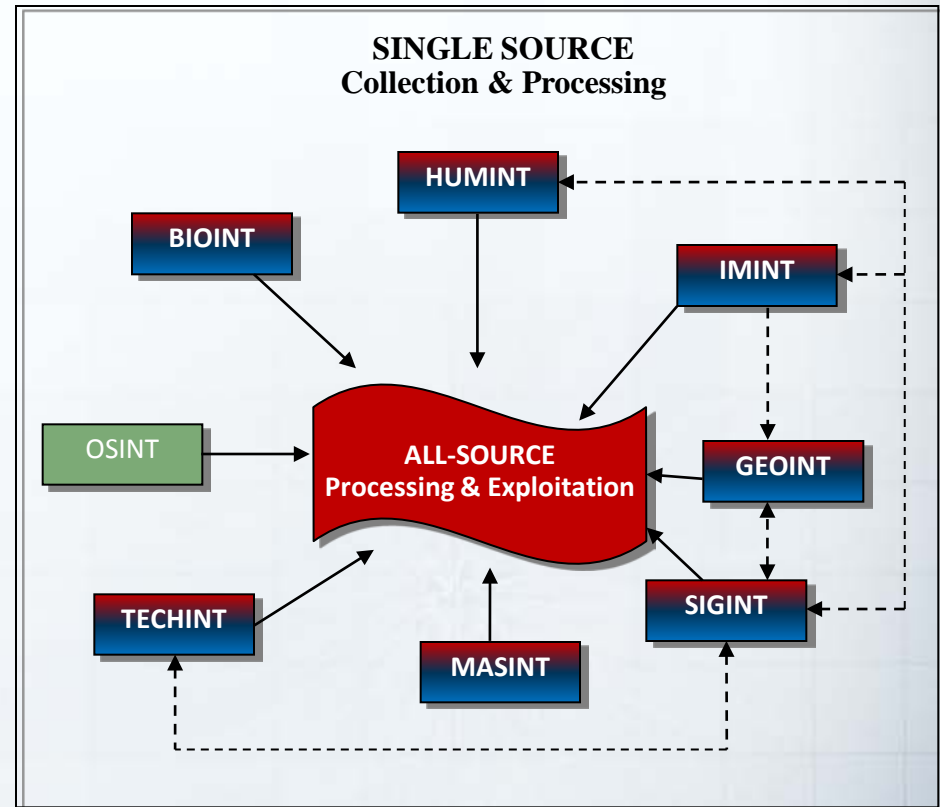
Army/Intelligence context

- Increasing amount of data/information to be managed
 - Stored in distributed , stove-piped systems
- Multi-sources: SIGINT, IMINT/GeoINT, HUMINT, OSINT, etc.
- Various formats: sensor data, multimedia (text, images, audio, video)
 - Hard/soft, structured/unstructured
- Dynamic real-time data vs static/historical data
- Information overload

Objectives

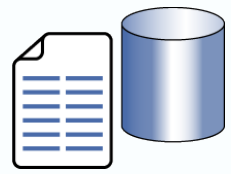
- Make data accessible, discoverable, integrated and disseminated
- Integrated approach for information management and integration
- User-oriented services for enhanced access and exploitation
 - **ASIM** (All-Source Information Integration and Management) R&D Project

All-source intelligence is both a separate intelligence discipline and the name of the process used to produce intelligence from multiple intelligence or information sources. (*US Army doctrine for intelligence FM-2.0*).



Single-Source / All-Source Intelligence Production

SIGINT



HUMINT



GeoINT



IMINT

MASINT



OSINT



...



Filter, index, organize

Analyze, fuse, synthesize

Evaluate

Disseminate

Collection

Processing

All-Source Information Management

- **Multiple sources and formats** (structured / unstructured)
 - Better describe, characterize information (metadata)
 - Make information accessible, discoverable, retrievable, exploitable
 - Automated metadata tagging
 - Facilitate information correlation among multiple sources

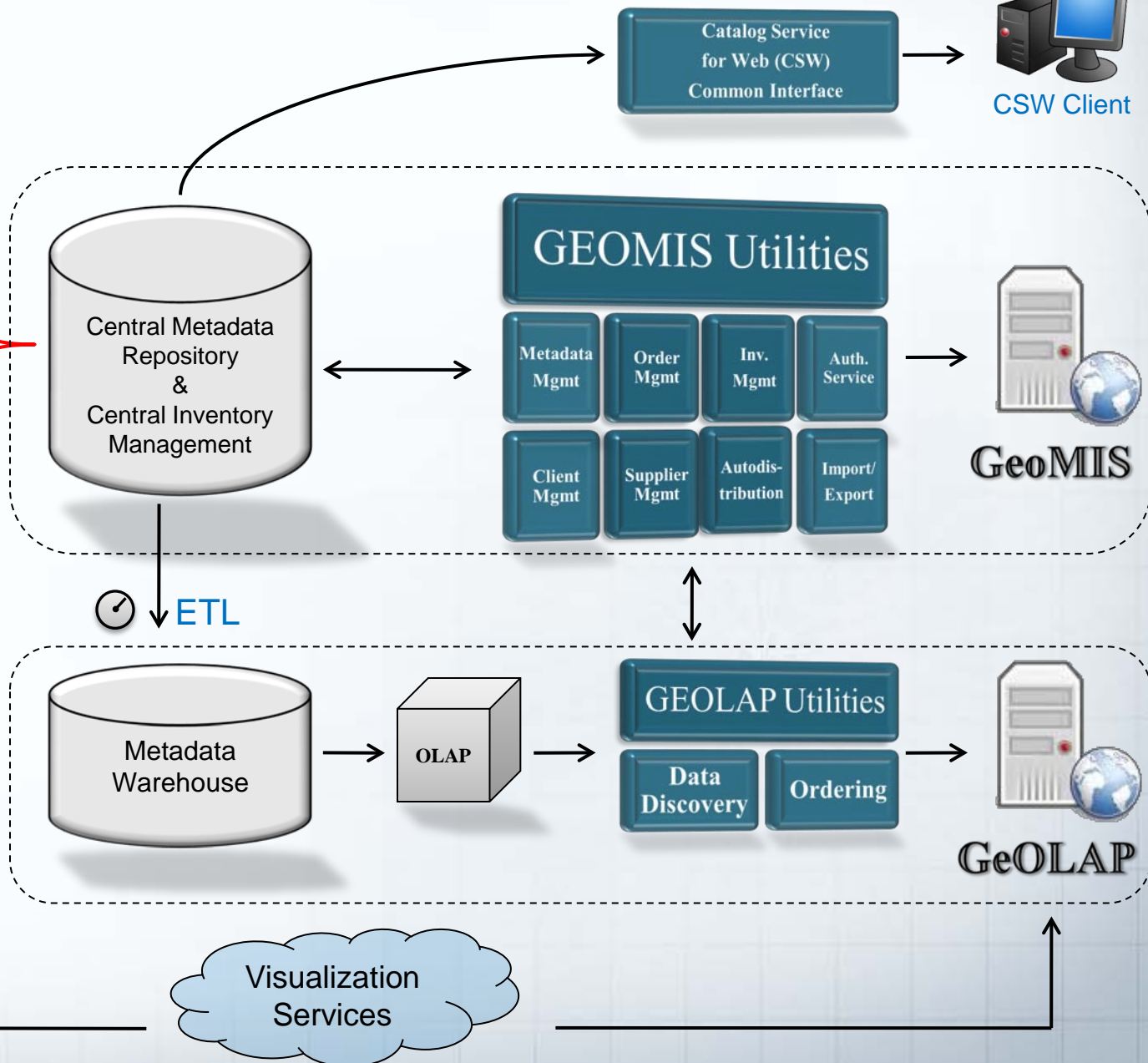
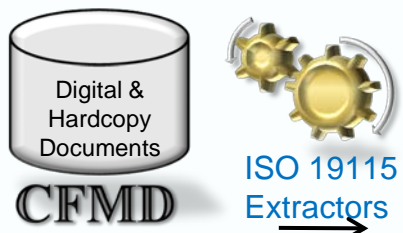
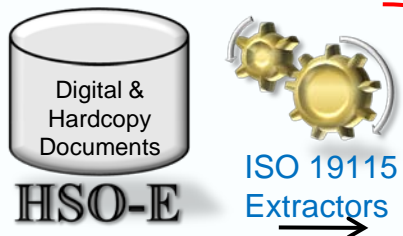
- **Exploitation of Standards**
 - Metadata
 - Information content (what, who, where), context, administrative
 - ISO, NATO, OGC (geospatial)
 - Ontologies (common terminology)
 - Formal semantics of objects
 - Integration, mapping

Geospatial Data Management

- **GeoMIS** (Geospatial Metadata and Inventory System)
 - Backbone infrastructure in the Mapping and Charting Establishment of DND to centrally manage all geospatial products
 - Manage large geospatial inventories through the use of metadata standards (e.g. ISO 19115).
 - Export to ISO 19139 XML, to KML
- **GeOLAP** (Spatial OLAP Data Discovery)
 - Provide a wide-audience online access to geospatial data holdings of DND through a map-based application.
 - Connected to GeoMIS, provide online visualization and ordering capabilities.
- **Extensions:**
 - Metadata exposed through OGC Catalog for the Web (CSW) implementation
 - Search and filtering capabilities

GEOLAP / GEOMIS

MCE Data Centers



All-Source Information/Intelligence Integration

- Hard/soft data correlation in a All-Source context
 - Standardized annotation of information elements
 - Facilitate all-source information correlation
- Integration of metadata standards within a federated environment
 - Metadata developed by COIs
 - Metadata mapping across domains
 - Exploitation of domain ontologies (or upper-ontology)
 - Metadata (RDF), ontologies (OWL)
 - E.g. Mapping US DDMS to SUMO (Veres & Ng)

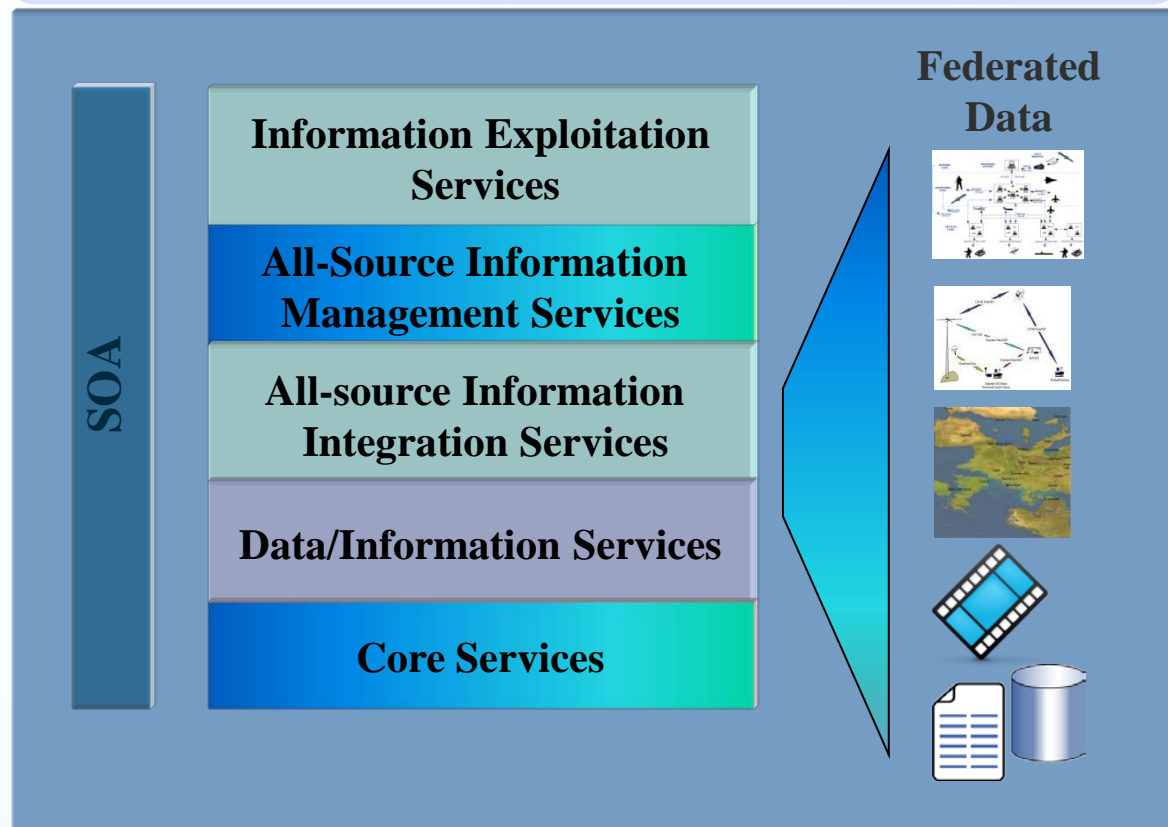
All-Source Information Management Framework

Objective: Provide a user-centric All-Source/Multi-int IM environment

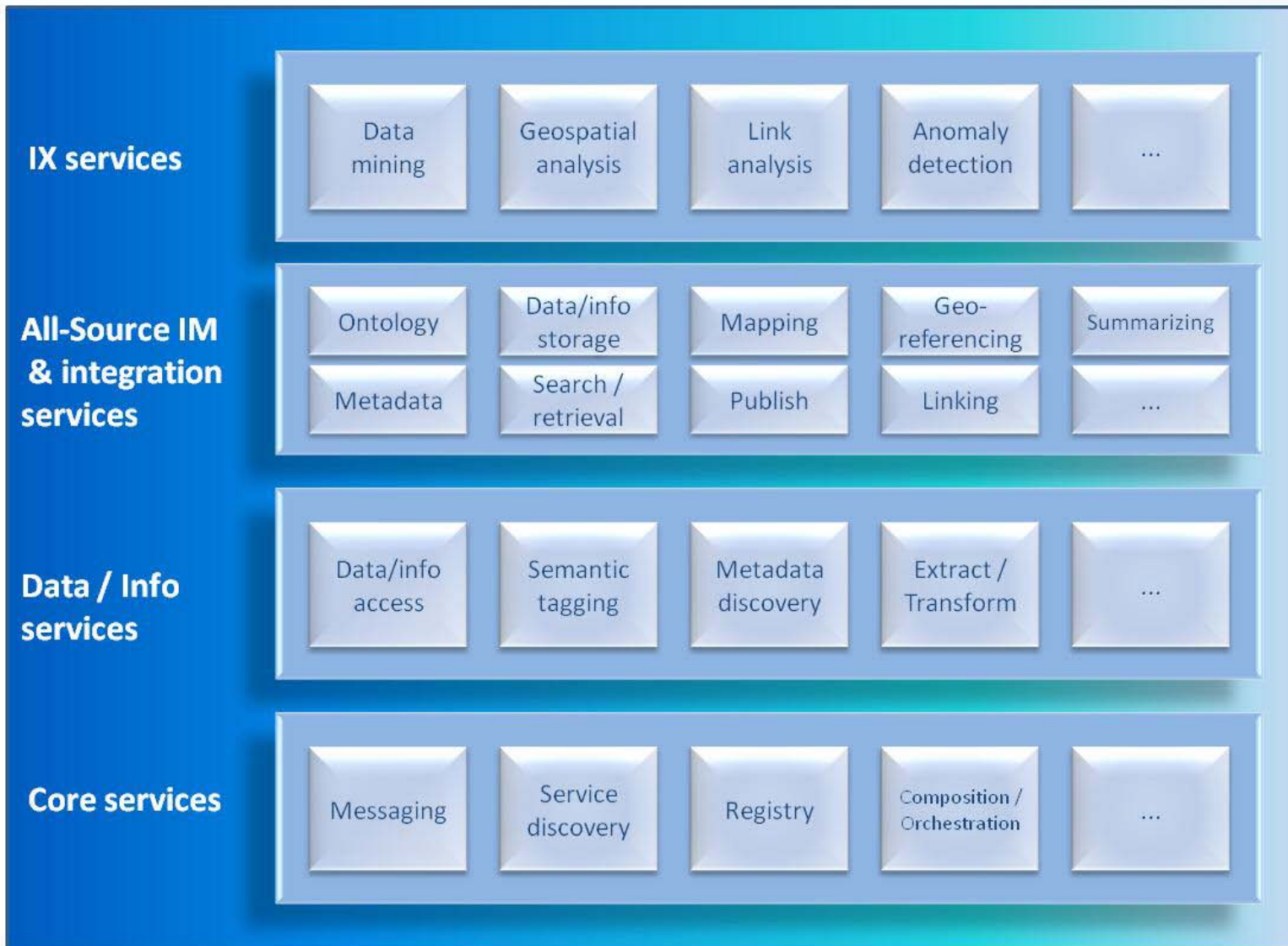
- Service-based IM environment
 - Open and flexible architecture
 - Exposure of loosely-coupled distributed services
 - Focus on IM and integration
 - Data/information access
 - Discovery, search and retrieval, notification
 - Mapping, correlation, synthesis
 - Rely on Core services
 - Support of Information Exploitation (IX) services

ASIM Framework

SOA Framework for All-Source Information Management and Integration



ASIM Categories of services



Conclusions and perspectives

- Information Management Challenges in a All-Source context
 - Enhance description of intelligence information from multiple sources
 - Governance required for enhanced IM and information sharing
- Enabling Technology
 - Service-based IM Framework
 - Maximize use of open standards and semantic technologies
- Future work
 - Refine the framework for All-Source information Management
 - Cross-COI interoperability
 - Handling of large data sets
 - Align with MAJIIC-2 (multi-int perspective)
 - Related initiatives (OGC Fusion Standards Study)



POC: Anne-Claire.Boury-Brisset@drdc-rddc.gc.ca